



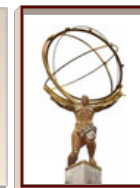
Tier 2 Storage Requirements

Kaushik De

Facilities Meeting, BNL

June 11, 2008

Introduction



- ❑ We need to estimate Tier 2 storage needs for data taking phase of ATLAS operations
- ❑ Complex exercise – fundamentals are changing daily
- ❑ But we need to start process, to plan for purchases
- ❑ First proposal here:
 - ❑ Estimate of what we need for 2008 run (~till end of the year)
 - ❑ Very preliminary – needs to be iterated

Assumptions



❑ Space tokens at Tier 2's:

- ❑ PROD/PRODDISK/PRODUCTION – transient, used by Panda
- ❑ MCDISK – physics AOD's and DPD's, centrally subscribed
- ❑ DATADISK – AOD's and DPD's from collider data
- ❑ USER/USERDISK – user (derived) data storage, from pathena
- ❑ GROUP/GROUPDISK – physics group DPD's

❑ Rate assumptions

- ❑ MC production rate to keep all CPU's busy
- ❑ 60 days of data from LHC
- ❑ 100% AOD+D1PD kept at each Tier 2
- ❑ 50% of MC AOD+D1PD kept at each Tier 2
- ❑ Each kSi2k produces 100GB of user data (integrated for 2008)
- ❑ Unresolved issue - transient vs permanent user data (Tier 1 issue?)

Requirements Analysis



❑ PRODUCTION

- ❑ Range 19-39 TB (based on CPU pledge)
- ❑ Assume 40 TB if we choose uniform requirement for each Tier 2

❑ MCDISK

- ❑ 66 TB (assuming 50% of all AOD+D1PD at each Tier 2)

❑ DATADISK

- ❑ 168 TB (100% of AOD and D1PD) – based on Jim's numbers

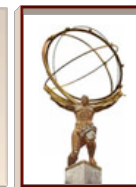
❑ GROUP

- ❑ 10 TB (just a guess)

❑ USER

- ❑ Range 17-35 TB
- ❑ Recommendation: 35 TB at each Tier 2

US T2 Status



	Required (all 5 tokens)	WLCG MoU Pledge 2008	Difference	US Regional Quota?
GLT2	319 TB	322 TB	3 TB	64 TB
NET2	319 TB	282 TB	-37 TB	64 TB
MWT2	319 TB	244 TB	-75 TB	64 TB
SWT2	319 TB	256 TB	-63 TB	64 TB
WT2	319 TB	462 TB	143 TB	64 TB

Comparisons with Typical EGEE Tier 2



□ Typical EGEE Tier 2:

- ~20% of AOD's
- Smaller user allocation
- Average 538 kSI2k (less than 50% compared to US Tier2's)

□ Caveats:

- The typical EGEE T2 is my estimate (same data and MC rate assumptions as US T2)
- I will share with Kors – but note that distributions among Tokens are slightly different than Kors' numbers last week, though total is same (100 TB)

	Typical EGEE T2	US T2
PROD	15 TB	40 TB
MCDISK	26 TB	66 TB
DATADISK	34 TB	168 TB
USER	15 TB	35 TB
GROUP	10 TB	10 TB

Summary



- ❑ Warning: these numbers are very preliminary
- ❑ Caveat: only for internal use and may change any time
- ❑ If we get 60 days of data this year, some US Tier 2's may fall short of 2008 pledges – new procurements should take this into account
- ❑ I am assuming that we can start with (relatively) clean slate (old data will be removed before 10 TeV production starts)
- ❑ In general, US data model works
- ❑ Big uncertainties: USER and GROUP
- ❑ Stay tuned – these number will certainly get revised
- ❑ Stay tuned – for 2009 numbers